Statistics Rules

- Pearson's correlation coefficient:

$$r = \frac{n\Sigma xy - \Sigma x \times \Sigma y}{\sqrt{n\Sigma x^2 - (\Sigma x)^2}\ \sqrt{n\Sigma y^2 - (\Sigma y)^2}}$$

- Spearman's rank correlation Coefficient:

$$r = 1 - \frac{6\Sigma D^2}{n(n^2 - 1)}$$

Regression line equation:

$$\hat{y} = a + b\,x$$

where b is the regression coefficient of y on x.

$$, b = \frac{n\Sigma xy - \Sigma x \times \Sigma y}{n\Sigma x^2 - (\Sigma x)^2} \quad , a = \frac{\Sigma y - b\Sigma x}{n}$$

- The regression line equation of y on x is used for predicting the value of y if the value of x is Known also in identifying the error which can be identified by the relation

Error = | table value – the value satisfying the equation |

The correlation is a relation between 2 variables or more.

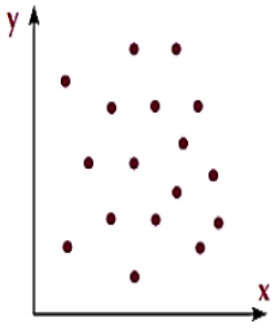


Figure 1
There is no correlation

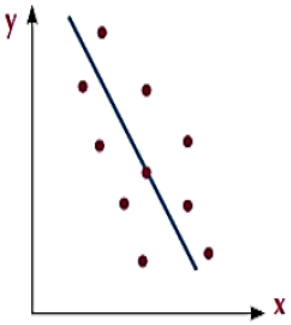


Figure 2
There is an inverse linear correlation

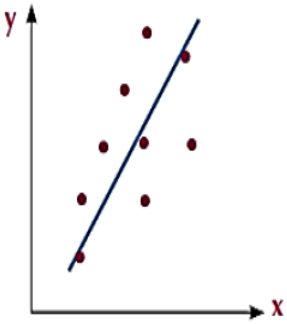
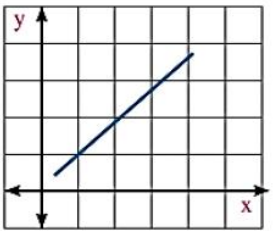
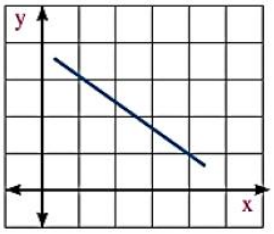


Figure 3
There is a direct linear correlation



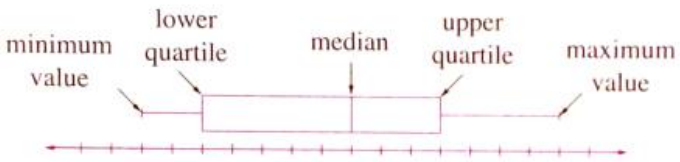
perfect direct correlation



perfect inverse correlation

D.M.RAZK موقع الدكتور محمد رزق معلم الكيمياء التعليمي

Representation by using box and whisker plot



$$\text{Semi - interquartile range} = \frac{\text{upper quartile} - \text{lower quartile}}{2}$$

Range – Mean – Mode – Median

- The range of set of values = The maximum value - The minimum value
- The arithmetic mean of set of values = $\frac{\text{The sum of all values}}{\text{The number of these values}}$
- The mode : It is the most frequently occuring value.
- The median : It is the value that exists in the middle of the values after arranging them.

probability:

- If A and B two events, then:

$P(A \cup B) = B(A) + P(B) - P(A \cap B)$

$P(A - B) = P(A) - P(A \cap B)$

$P(B') = 1 - P(B)$

- If A and B are two mutually exclusive events, then:

$A \cap B = \phi$, $P(A \cap B) = 0$

- Conditional probability:

If S the sample space of a random experiment, A and B are two events of this space, then the probability of occurring the event A in condition of occurring the event B

$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$ where $P(B) > 0$

D.M.RAZK موقع الدكتور محمد رزق معلم الكيمياء التعليمي

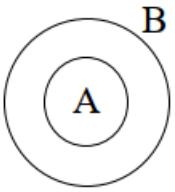
5) $P(A)+P(A')=1$

$\therefore P(A')=1-P(A)$ *also* $P(A)=1-P(A')$

6) $P(A' \cap B')=P(A \cup B)'=1-P(A \cup B)$

8) $P(A \cap B')=P(A-B)$

9) If $A \subset B$, then $A \cup B=B$, $A \cap B=A$, $A-B=\phi$



$P(A \cap B)=P(A)$, $P(A \cup B)=P(B)$

$P(A-B)=0$, $P(B-A)=P(B)-P(A)$

- The two independent events:

It is said that the two events A and B are independent events if and only if

$P(A \cap B) = P(A) \times P(B)$

Random variables:

- The probability distribution of a discrete random variable

x_r	x_1	x_2	x_3	.....	x_n
$f(x_r)$	$f(x_1)$	$f(x_2)$	$f(x_3)$	.....	$f(x_n)$

1- $f(x_r) \geqslant 0$ for each $r = 1, 2, 3, ..., n$

2- $f(x_1) + f(x_2) + f(x_3) + ... + f(x_n) = 1$

- Expectation (Mean) $(\mu) = \sum_{r=1}^{n} x_r \times f(x_r)$

- Variance : $\sigma^2 = \sum_{r=1}^{n} x^2_r \times f(x_r) - \mu^2$

- Standard deviation: σ = Positive Square root of the variance

---

- Coefficient of variation $= \frac{\sigma}{\mu} \times 100\ \%$

Geometric probability distribution function

If $X \sim G(S)$ then

$$P(X = n) = S(1 - S)^{n-1} \qquad n = 1, 2, 3, ...$$

Where S is the probability of success, n is the number of trails until reach to the first success

D.M.RAZK موقع الدكتور محمد رزق معلم الكيمياء التعليمي

Expectation and variance of geometric distribution

Mean (expectation)	μ	$= \frac{1}{S}$
Variance	σ^2	$= \frac{1 - S}{S^2}$
Standard deviation	σ	= the positive square root of the variance

Probability distribution of a binomial random variable

If $X \sim B(n, S)$ then

$P(X = r) = {}^{n}C_{r} \times S^{r} \times (1 - S)^{n - r}$, $r = 1, 2, 3, \ldots n$

n: number of trails in the experiment

S:probability of success in each trail

r :is the number of trails (required number)

Expectation and variance of binomial distribution

If X is a random variable $X \sim B(n, S)$

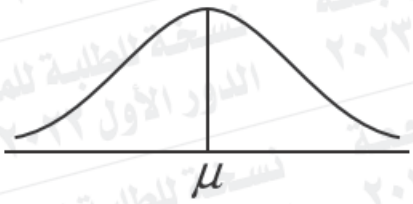
Mean (expectation) $\mu = n \times S$

Variance $\sigma^2 = n \times S \times (1 - S)$

Standard deviation σ = the positive square root of the variance

Where n is the number of trails in the experiment, S is the probability of success in each trail

Normal distribution:



1- The area of the region under the normal curve and above the x axis equals 1.

2- From symmetry, we find that the straight line $x = \mu$ divides the area under the curve and above the x - axis into two regions, the area of each region = 0.5

- To convert a normal distribution into the standard normal distribution we use the relation:

$$Z = \frac{x - \mu}{\sigma}$$

D.M.RAZK موقع الدكتور محمد رزق معلم الكيمياء التعليمي

Estimation Error

When a sample is used to estimate the population mean, the error in the estimate is represented by the symbol E at a confidence level of 1 - α, which is determined by the following relationship: $E = \frac{\sigma}{\sqrt{n}} \times Z_{\frac{\alpha}{2}}$

Where σ is the standard deviation and n is the sample size

Confidence interval for mean population μ

If a random sample of size n is taken from a population that follows a normal distribution with mean μ and variance σ^2

Then $\mu \in]\ \overline{X} - E\ ,\ \overline{X} + E\ [$

D.M.RAZK موقع الدكتور محمد رزق معلم الكيمياء التعليمي

موقع الدكتور
محمد رزق
التعليمي

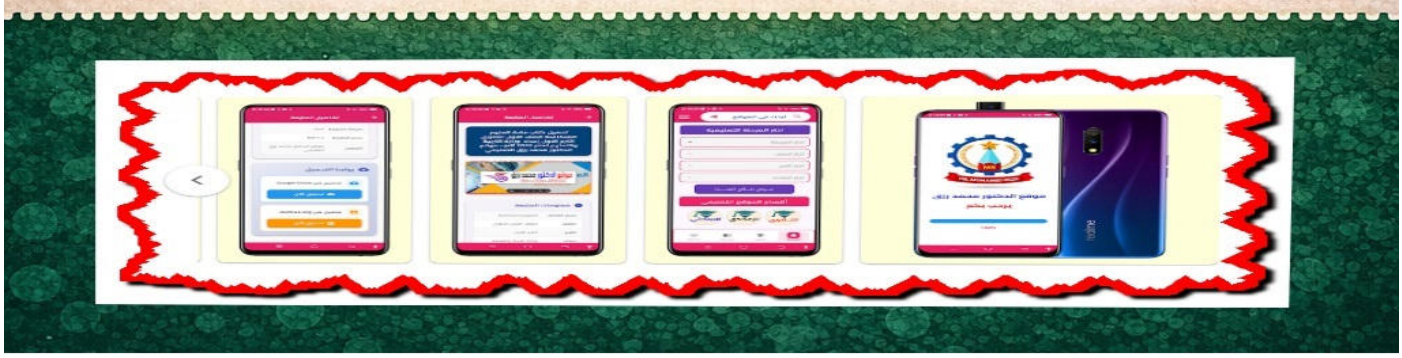
امسح الباركود بالموبيل لتثبيت
التطبيق وتحميل كل الملازم مجانا وبسرعه

MR

حمل التطبيق الأن

GET IT ON
Google Play

ملازم ابتدائي | ملازم اعدادي | ملازم ثانـــوي | كورسات ونتائج امتحانات



موقع الدكتور محمد رزق
Dr.MoRizK